[This Page Intentionally Left Blank]

7.0 RELIABILITY OF THE 3T3 AND NHK NRU TEST METHODS

This section discusses the reliability of the 3T3 and NHK NRU test methods. Reliability is the degree to which a test method can be performed reproducibly within and among laboratories over time (ICCVAM 2003). It is assessed by calculating intra- and inter-laboratory reproducibility and repeatability. Reproducibility is the consistency of individual test results obtained in a single laboratory (intralaboratory reproducibility) or in different laboratories (interlaboratory reproducibility) using the same protocol and test samples. Repeatability, usually applied to results within a laboratory, is the closeness of agreement between test results obtained within a single laboratory when the procedure is performed on the same substance under identical conditions within a given time. The NICEATM/ECVAM study was not designed to assess intralaboratory repeatability.

For the NICEATM/ECVAM validation study, reliability was assessed by determining both intra- and inter-laboratory reproducibility. Intralaboratory reproducibility is the agreement of results produced when qualified people within the same laboratory perform the test method using the same test protocol at different times (ICCVAM 2003). Interlaboratory reproducibility is the agreement of results from different qualified laboratories using the same protocol and reference substances. Interlaboratory reproducibility indicates the extent to which a test method can be successfully transferred among laboratories.

Intra- and inter-laboratory reproducibility of the 3T3 and NHK NRU test methods were determined using ANOVA and CV analysis as discussed in **Section 5.3.3 (**see **Sections 7.2.1** and **7.2.2**). Interlaboratory reproducibility of the 3T3 and NHK NRU test methods was also assessed by comparing the laboratory-specific IC_{50}-LD_{50} regressions (from **Table 6-1**) to one another for each test method (see **Section 7.2.3**) and by evaluating laboratory concordance for the GHS acute oral toxicity category predictions reported in **Sections 6.3.1** through **6.3.3** (see **Section 7.2.4**). Laboratory concordance for the solvent selection process using the solubility protocol (described in **Section 2.9**) is provided in **Section 7.4**.

7.1 Substances Used to Determine the Reliability of the 3T3 and NHK NRU Test Methods

The SMT intended to use the IC_{50} results of all 72 reference substances identified for testing in **Table 3-2** to determine the reliability of the 3T3 and NHK NRU test methods. Unfortunately, IC_{50} results for all substances could not be obtained in all the laboratories. **Table 7-1** shows the substances that failed to yield sufficient cytotoxicity for the calculation of an IC_{50} and the number of substances left to determine intralaboratory reproducibility. The laboratories failed to obtain IC_{50} results for three to five substances in the 3T3 NRU test method and two to three substances with the NHK NRU test method.

For the 3T3 NRU test method, no laboratory achieved sufficient cytotoxicity to obtain IC_{50} values for carbon tetrachloride or methanol and only one laboratory obtained IC_{50} results for lithium carbonate and xylene. Thus, interlaboratory reproducibility for the 3T3 NRU test method was assessed using the remaining 68 reference substances. For the NHK NRU test method, no laboratory obtained IC_{50} values for carbon tetrachloride and only one laboratory achieved IC_{50} results for xylene and 1,1,1-trichloroethane. Interlaboratory reproducibility for the NHK NRU test method was assessed using the IC_{50} results for the remaining 69 reference substances.

Despite the fact that IC_{50} values were not obtained by all the laboratories for all reference substances, **Table 7-2** shows that the complete range of LD_{50} responses, as defined by the GHS classification for acute oral toxicity in **Table 3-1**, was covered by the remaining substances. The IC_{50} values also covered a wide range of responses (see **Table 7-3**). IC_{50} values for the 3T3 NRU test method ranged from 0.005 to 38,878 μg/mL. IC_{50} values for the NHK NRU test method covered a larger range, from 0.00005 to 49,800 μg/mL.

Table 7-1 Reference Substances That Failed to Yield IC_{50} Values[1] And Number of Reference Substances Available for Intralaboratory Reproducibility Analyses

Laboratory	3T3 NRU Test Method		NHK NRU Test Method	
	Reference Substances Lacking IC_{50} Results	N[2]	Reference Substances Lacking IC_{50} Results	N[2]
ECBC	Carbon tetrachloride Methanol Xylene	69	Carbon tetrachloride Methanol Xylene	69
FAL	Carbon tetrachloride Gibberellic acid Lithium carbonate Methanol Xylene	67	1,1,1-Trichloroethane Carbon tetrachloride Xylene	69
IIVS	Carbon tetrachloride Lithium carbonate Methanol	69	1,1,1-Trichloroethane Carbon tetrachloride	70

[1]Due to insufficient cytotoxicity.
[2]Number of substances available for intralaboratory reproducibility analyses.

Table 7-2 Number of Reference Substances Tested vs Number of Reference Substances Yielding IC_{50} Values in Each GHS Toxicity Category[1] for Two Sets of LD_{50} Values

GHS Category[1] (LD_{50} in mg/kg)	Initial Oral LD_{50}[2]	Reference Oral LD_{50}[3]	Results from 3T3 NRU Test Method		Results from NHK NRU Test Method	
			Initial Oral LD_{50}[2]	Reference Oral LD_{50}[3]	Initial Oral LD_{50}[2]	Reference Oral LD_{50}[3]
$LD_{50} \leq 5$	12	7	12	7	12	7
$5 < LD_{50} \leq 50$	12	12	12	12	12	12
$50 < LD_{50} \leq 300$	12	12	12	12	12	12
$300 < LD_{50} \leq 2000$	12	16	11	15	12	16
$2000 < LD_{50} \leq 5000$	12	12	10	10	10	10
$LD_{50} > 5000$	12	13	11	12	11	12

[1]GHS-Globally Harmonized System of Classification and Labelling of Chemicals (UN 2005).
[2]Number of reference substances that yielded an IC_{50} value in at least one laboratory based on initial oral LD_{50} in **Table 3-2**. Initial oral LD_{50} values, used during the reference substance selection process, were those used by the Registry of Cytotoxicity (RC) (from 1983/84 RTECS®) when applicable. The RC is a database of acute oral LD_{50} values for rats and mice obtained from RTECS® and IC_{50} values from *in vitro* cytotoxicity assays using multiple cell lines and cytotoxicity endpoints for chemicals with known molecular weights (Halle 1998). Values for reference substances not included in the RC came from HSDB or RTECS®.
[3]Number of reference substances that yielded an IC_{50} value in at least one laboratory based on reference oral LD_{50} in **Table 4-2**. Reference oral LD_{50} values from rats and mice were derived after evaluating LD_{50} values located through literature searches and references from toxicity databases such as RTECS®.

7.2 Reproducibility Analyses for the 3T3 and NHK NRU Test Methods

Reproducibility of the 3T3 and NHK NRU test methods were performed using ANOVA and CV as described in **Section 5.3.3**. **Table 7-3** reports the results of these analyses for each reference substance and test method.

7.2.1 ANOVA Results for the 3T3 and NHK NRU Test Methods

ANOVA was performed as discussed in Section **5.3.3**. Since the sample sizes from this study were small, usually three observations per laboratory, the ANOVA results may be misleading. There may be some differences that are statistically significant only because there are too few observations within the laboratories to adequately characterize the variability, and/or the within-laboratory variability estimate is small.

Differences Among the Laboratories for the 3T3 NRU Test Method

The ANOVA results in **Table 7-3** indicate that there were statistically significant ($p < 0.01$) differences among the laboratories for 26 reference substances. These chemicals are listed in **Table 7-4** along with columns showing the laboratory statistically significantly differing from the other two laboratories (as indicated by the contrast results). Since significant laboratory differences may be produced by insolubility or volatility, **Table 7-4** also indicates whether any laboratory reported insolubility or volatility during conduct of the test. Insolubility was suggested by the presence of precipitates in either the stock solutions or in cell culture. Volatility was identified by the use of plate sealers to contain volatile contamination of lower concentration wells by higher concentrations. Insolubility and volatility were reported for only nine of the 26 chemicals.

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
Acetaminophen	50.1		28	1.70	0.171		526		13	2.72	0.181	
ECBC	40.8	22		1.61		NA	558	15		2.75		NA
FAL	66.2	35		1.82		NA	447	19		2.65		NA
IIVS	43.4	26		1.64		NA	571	14		2.76		NA
Acetonitrile	8484		21	3.93	0.553		10104		8	4.00	0.9641	
ECBC	6433	2		3.81		NA	10868	72		4.04		NA
FAL	9690	58		3.99		NA	10153	19		4.01		NA
IIVS	9330	13		3.97		NA	9290	4		3.97		NA
Acetylsalicylic acid	760		56	2.88	<0.001		613		15	2.79	0.060	
ECBC	646	10		2.81		0.581	631	3		2.80		NA
FAL	1234	24		3.09		<0.001	694	14		2.84		NA
IIVS	401	16		2.60		<0.001	514	15		2.71		NA
5-Aminosalicylic acid	1698		19	3.23	0.054		52.3		47	1.72	0.044	
ECBC	1467	14		3.17		0.092	29.9	22		1.48		0.025
FAL	2070	16		3.32		0.021	78.2	54		1.89		0.033
IIVS	1557	12		3.19		0.312	48.8	16		1.69		0.832
Aminopterin	0.007		54	-2.14	0.036		682		27	2.83	0.0250	
ECBC	0.005	20		-2.28		0.216	889	20		2.95		0.017
FAL	0.012	46		-1.93		0.013	545	8		2.74		0.041
IIVS	0.005	23		-2.33		0.079	611	12		2.79		0.345
Amitriptyline HCl	7.23		14	0.86	0.348		9.76		19	0.99	0.365	
ECBC	6.03	23		0.78		0.163	10.8	31		1.03		NA
FAL	7.86	28		0.90		0.469	7.57	72		0.88		NA
IIVS	7.81	18		0.89		0.445	10.9	10		1.04		NA
Arsenic trioxide	2.51		61	0.40	0.004		10.4		91	1.02	<0.001	
ECBC	2.41	33		0.38		0.527	7.77	33		0.89		0.694
FAL	1.04	7		0.02		0.002	2.55	75		0.41		<0.001
IIVS	4.09	52		0.61		0.006	20.9	31		1.32		0.0006
Atropine sulfate	85.6		49	1.93	0.049		91.9		13	1.96	0.9881	
ECBC	54.1	55		1.73		0.046	85.4	12		1.93		0.8903
FAL	133	31		2.12		0.024	104	85		2.02		0.9069
IIVS	70.0	8		1.85		0.641	83.2	25		1.92		0.9832
Boric acid	2228		69	3.35	0.010		473		8	2.67	0.9306	
ECBC	1497	32		3.18		0.189	440	31		2.64		0.9692
FAL	3987	17		3.60		0.004	517	73		2.71		0.7391
IIVS	1202	48		3.08		0.021	464	2		2.67		0.7680
Busulfan	135		119	2.13	0.002		278		11	2.44	0.659	
ECBC	40.0	48		1.60		0.012	253	27		2.40		NA
FAL	321	56		2.51		< 0.001	268	72		2.43		NA
IIVS	43.7	4		1.64		0.033	313	12		2.50		NA
Cadmium chloride	0.565		39	-0.25	0.124		1.98		10	0.30	0.733	
ECBC	0.480	14		-0.32		NA	2.20	37		0.34		NA

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
FAL	0.400	32		-0.40		NA	1.88	65		0.27		NA
IIVS	0.817	53		-0.09		NA	1.86	8		0.27		NA
Caffeine	161		18	2.21	0.481		661		21	2.82	0.296	
ECBC	133	10		2.12		NA	817	31		2.91		NA
FAL	157	52		2.20		NA	591	32		2.77		NA
IIVS	191	7.5		2.28		NA	574	1		2.76		NA
Carbamazepine	109		35	2.04	0.049		128		85	2.11	0.432	
ECBC	83.0	14		1.92		NA	66.1	13		1.82		NA
FAL	152	37		2.18		NA	253	129		2.40		NA
IIVS	91.8	12		1.96		NA	63.9	8		1.81		NA
Carbon tetrachloride	NA		NA	NA	NA		NA		NA	NA	NA	
ECBC	NA	NA		NA		NA	NA	NA		NA		NA
FAL	NA	NA		NA		NA	NA	NA		NA		NA
IIVS	NA	NA		NA		NA	NA	NA		NA		NA
Chloral hydrate	187		25	2.27	0.004		137		17	2.14	0.302	
ECBC	151	10		2.18		0.008	140	24		2.15		NA
FAL	241	10		2.38		0.002	159	32		2.20		NA
IIVS	170	12		2.23		0.181	112	2		2.05		NA
Chloramphenicol	161		67	2.21	<0.001		366		13	2.56	0.750	
ECBC	55.3	22		1.74		<0.001	318	45		2.50		NA
FAL	273	30		2.44		0.001	414	44		2.62		NA
IIVS	156	18		2.19		0.165	367	22		2.56		NA
Citric acid	829		41	2.92	0.002		424		25	2.63	0.006	
ECBC	473	29		2.68		0.001	526	16		2.72		0.009
FAL	1148	13		3.06		0.003	312	17		2.49		0.002
IIVS	865	19		2.94		0.298	433	5		2.64		0.483
Colchicine	0.047		85	-1.33	0.001		0.007		22	-2.16	0.174	
ECBC	0.020	11		-1.70		0.0028	0.005	46		-2.28		NA
FAL	0.093	45		-1.03		0.0005	0.008	10		-2.12		NA
IIVS	0.028	1		-1.55		0.0914	0.008	21		-2.09		NA
Cupric sulfate pentahydrate	70.6		85	1.85	<0.001		197		4	2.29	0.374	
ECBC	82.7	4		1.92		0.001	190	10		2.28		NA
FAL	123	44		2.09		<0.001	195	6		2.29		NA
IIVS	5.70	31		0.76		<0.001	207	3		2.32		NA
Cycloheximide	0.293		104	-0.53	0.021		0.082		43	-1.09	0.302	
ECBC	0.125	45		-0.90		0.118	0.053	22		-1.28		NA
FAL	0.647	70		-0.19		0.007	0.120	78		-0.92		NA
IIVS	0.109	23		-0.96		0.076	0.071	19		-1.15		NA
Dibutyl phthalate	78.3		124	1.89	< 0.001		32.6		41	1.51	0.408	
ECBC	23.5	17		1.37		0.012	28.3	27		1.45		NA

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
FAL	191	50		2.28		<0.001	47.4	73		1.68		NA
IIVS	20.7	7		1.32		0.005	22.0	6		1.34		NA
Dichlorvos	20.3		57	1.31	0.002		11.1		20	1.05	0.181	
ECBC	9.80	35		0.99		0.001	8.56	27		0.93		NA
FAL	32.8	6		1.52		0.002	12.4	30		1.09		NA
IIVS	18.3	11		1.26		0.823	12.2	3		1.09		NA
Diethyl phthalate	113		28	2.05	0.127		145		44	2.16	0.049	
ECBC	85.5	34		1.93		0.092	174	8		2.24		0.196
FAL	147	26		2.17		0.070	71.5	94		1.85		0.018
IIVS	106	24		2.03		0.846	189	18		2.28		0.127
Digoxin	520		62	2.72	0.043		0.00314		88	-2.50	<0.001	
ECBC	351	39		2.54		0.167	0.00538	13		-2.27		<0.001
FAL	892	36		2.95		0.017	0.00005	36		-4.29		<0.001
IIVS	317	21		2.50		0.144	0.00398	7		-2.40		<0.001
Dimethylformamide	5242		6	3.72	0.296		7856		19	3.90	<0.001	
ECBC	5343	10		3.73		NA	9353	2		3.97		<0.001
FAL	5483	9		3.74		NA	7817	1		3.89		0.508
IIVS	4900	4		3.69		NA	6397	3		3.81		<0.001
Diquat dibromide monohydrate	15.1		120	1.18	0.017		4.73		37	0.67	0.217	
ECBC	3.90	23		0.59		0.040	3.59	23		0.56		NA
FAL	36.1	98		1.56		0.006	6.77	55		0.83		NA
IIVS	5.40	25		0.73		0.190	3.84	8		0.58		NA
Disulfoton	98.6		55	1.99	0.003		378		99	2.58	<0.001	
ECBC	137	55		2.14		NA	140	19		2.15		0.002
FAL	NA	NA		NA		NA	808	26		2.91		<0.001
IIVS	60.4	87		1.78		NA	186	32		2.27		0.018
Endosulfan	8.02		78	0.90	0.046		2.35		43	0.37	0.029	
ECBC	5.30	57		0.72		0.447	3.44	17		0.54		0.020
FAL	15.2	78		1.18		0.018	1.42	50		0.15		0.018
IIVS	3.60	42		0.56		0.080	2.19	20		0.34		0.927
Epinephrine bitartrate	59.4		12	1.77	0.048		90.6		24	1.96	0.119	
ECBC	51.5	12		1.71		0.018	115	9		2.06		NA
FAL	63.4	11		1.80		0.165	81.7	35		1.91		NA
IIVS	63.4	3		1.80		0.149	75.0	16		1.88		NA
Ethanol	6731		23	3.83	0.075		10184		18	4.01	0.035	
ECBC	5360	33		3.73		NA	8290	5		3.92		0.019
FAL	8420	14		3.93		NA	12013	19		4.08		0.029
IIVS	6413	5		3.81		NA	10250	9		4.01		0.752
Ethylene glycol	25292		26	4.40	0.007		42600		15	4.63	0.063	
ECBC	18325	9		4.26		0.004	38000	12		4.58		NA
FAL	31650	24		4.50		0.010	49800	9		4.70		NA

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
IIVS	25900	12		4.41		0.505	40000	13		4.60		NA
Fenpropathrin	27.2		49	1.43	0.301		2.60		39	0.41	0.031	
ECBC	22.6	11		1.35		NA	3.73	27		0.57		0.013
FAL	42.4	63		1.63		NA	2.23	28		0.35		0.375
IIVS	16.7	12		1.22		NA	1.82	17		0.26		0.044
Gibberellic Acid	7842		3	3.89	0.621		2866		2	3.46	0.862	
ECBC	8027	11		3.90		NA	2850	14		3.45		NA
FAL	NA	NA		NA		NA	2940	9		3.47		NA
IIVS	7657	10		3.88		NA	2807	4		3.45		NA
Glutethimide	192		43	2.28	< 0.001		177		5	2.25	0.968	
ECBC	167	4		2.22		0.029	187	34		2.27		NA
FAL	284.3	7		2.45		<0.001	170	14		2.23		NA
IIVS	125.3	7		2.10		<0.001	176	16		2.24		NA
Glycerol	28904		33	4.46	0.846		27108		31	4.43	0.200	
ECBC	20000	15		4.30		NA	34267	45		4.53		NA
FAL	38878	73		4.59		NA	18023	46		4.26		NA
IIVS	27833	39		4.44		NA	29033	16		4.46		NA
Haloperidol	6.26		24	0.80	0.006		3.57		7	0.55	0.935	
ECBC	5.30	12		0.72		0.030	3.69	27		0.57		NA
FAL	8.00	8		0.90		0.002	3.72	49		0.57		NA
IIVS	5.50	12		0.74		0.061	3.29	35		0.52		NA
Hexachlorophene	4.48		27	0.65	0.174		0.031		41	-1.50	0.097	
ECBC	5.00	48		0.70		NA	0.027	16		-1.57		NA
FAL	5.30	33		0.72		NA	0.046	44		-1.34		NA
IIVS	3.10	9		0.49		NA	0.021	11		-1.67		NA
Lactic acid	3073		12	3.49	0.160		1308		1	3.12	0.904	
ECBC	2943	11		3.47		NA	1290	4		3.11		NA
FAL	3487	16		3.54		NA	1320	5		3.12		NA
IIVS	2790	9		3.45		NA	1313	11		3.12		NA
Lindane	161		58	2.21	0.066		19.3		20	1.29	0.203	
ECBC	125	95		2.10		NA	19.1	17		1.28		NA
FAL	266	36		2.43		NA	23.2	31		1.37		NA
IIVS	90.4	122		1.96		NA	15.6	15		1.19		NA
Lithium carbonate	NA		NA	NA	NA	NA	477		13	2.68	0.295	
ECBC	564	12		2.75		NA	411	29		2.61		NA
FAL	NA	NA		NA		NA	486	20		2.69		NA
IIVS	NA	NA		NA		NA	535	6		2.73		NA
Meprobamate	539		54	2.73	<0.001		516		61	2.71	0.027	
ECBC	353	14		2.55		0.001	761	15		2.88		0.0758
FAL	877	15		2.94		<0.001	163	116		2.21		0.0098
IIVS	386	2		2.59		0.005	624	14		2.80		0.1648

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
Mercury chloride	4.32		33	0.64	0.021		5.87		15	0.77	0.120	
ECBC	3.50	5		0.54		0.083	6.87	15		0.84		NA
FAL	6.00	31		0.78		0.008	5.40	19		0.73		NA
IIVS	3.50	3		0.54		0.110	5.35	2		0.73		NA
Methanol	NA		NA	NA	NA	NA	1616		42	3.21	0.007	
ECBC	NA	NA		NA		NA	NA	NA		NA		NA
FAL	NA			NA		NA	1133	19		3.05		NA
IIVS	NA			NA		NA	2100	11		3.32		NA
Nicotine	378		25	2.58	0.128		113		17	2.05	0.700	
ECBC	272	24		2.43		NA	94.3	26		1.97		NA
FAL	412	33		2.61		NA	134	59		2.13		NA
IIVS	450	12		2.65		NA	112	25		2.05		NA
Paraquat	23.3		8	1.37	1.000		66.1		40	1.82	0.047	
ECBC	21.3	34		1.33		NA	48.3	13		1.68		0.089
FAL	24.9	67		1.40		NA	96.6	39		1.98		0.018
IIVS	23.7	64		1.37		NA	53.4	10		1.73		0.279
Parathion	61.8		111	1.79	0.014		31.4		8	1.50	0.845	
ECBC	22.7	53		1.36		0.064	34.0	30		1.53		NA
FAL	141	70		2.15		0.005	31.2	38		1.49		NA
IIVS	22	22		1.34		0.081	29.0	29		1.46		NA
Phenobarbital	612		21	2.79	0.232		478		39	2.68	0.027	
ECBC	634	21		2.80		NA	693	26		2.84		0.010
FAL	726	35		2.86		NA	360	27		2.56		0.072
IIVS	476	23		2.68		NA	381	18		2.58		0.173
Phenol	70.9		41		0.011		77.7		22	1.89	0.094	
ECBC	50.2	22		1.70		0.022	59.1	36		1.77		NA
FAL	104	24		2.02		0.004	93.2	6		1.97		NA
IIVS	58.1	12		1.76		0.206	80.8	6		1.91		NA
Phenylthiourea	119		90	2.08	0.007		346		19	2.54	0.133	
ECBC	30.1	66		1.48		0.004	363	16		2.56		NA
FAL	239	28		2.38		0.006	401	21		2.60		NA
IIVS	89	25		1.95		0.718	272	26		2.44		NA
Physostigmine	28.8		30	1.46	0.149		172		22	2.24	0.623	
ECBC	28.2	53		1.45		NA	164	3		2.21		NA
FAL	37.8	5		1.58		NA	213	112		2.33		NA
IIVS	20.4	33		1.31		NA	139	6		2.14		NA
Potassium chloride	3635		7	3.56	0.846		2279		13	3.36	0.396	
ECBC	3352	14		3.53		NA	2560	17		3.41		NA
FAL	3842	31		3.58		NA	2287	28		3.36		NA
IIVS	3710	11		3.57		NA	1990	8		3.30		NA
Potassium cyanide	64.3		127	1.81	<0.001		45.1		86	1.65	0.340	
ECBC	15.3	25		1.18		0.001	29.3	24		1.47		NA

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
FAL	159	52		2.20		<0.001	89.0	112		1.95		NA
IIVS	18.9	5		1.28		0.006	16.9	13		1.23		NA
Procainamide HCl	443		11	2.65	0.007		1764		16	3.25	0.053	
ECBC	400	4		2.60		0.008	1480	14		3.17		NA
FAL	431	1		2.63		0.396	1787	12		3.25		NA
IIVS	497	8		2.70		0.003	2027	11		3.31		NA
2-Propanol	3563		23	3.55	0.001		5541		26	3.74	0.033	
ECBC	2610	9		3.42		< 0.001	5263	11		3.72		0.797
FAL	3970	4		3.60		0.004	4273	27		3.63		0.026
IIVS	4110	4		3.61		0.002	7087	7		3.85		0.018
Propranolol HCl	14.9		16	1.17	0.488		36.9		21	1.57	0.003	
ECBC	13.6	32		1.13		NA	38.27	12		1.58		0.325
FAL	13.5	51		1.13		NA	43.8	6		1.64		0.006
IIVS	17.6	21		1.25		NA	28.6	11		1.46		0.001
Propylparaben	29.9		64	1.48	0.001		16.8		16	1.23	0.066	
ECBC	20.9	16		1.32		0.045	18.1	13		1.26		NA
FAL	51.8	29		1.71		< 0.001	18.6	15		1.27		NA
IIVS	17.1	12		1.23		0.003	13.8	9		1.14		NA
Sodium arsenite	0.873		55	-0.06	0.028		0.532		44	-0.27	0.061	
ECBC	0.500	6		-0.30		0.032	0.790	32		-0.10		NA
FAL	1.40	57		0.15		0.012	0.336	56		-0.47		NA
IIVS	0.700	17		-0.15		0.478	0.470	14		-0.33		NA
Sodium chloride	4764		3	3.68	0.759		2724		51	3.44	0.045	
ECBC	4790	5		3.68		NA	3583	7		3.55		0.141
FAL	4625	13		3.67		NA	1118	124		3.05		0.017
IIVS	4877	9		3.69		NA	3470	9		3.54		0.161
Sodium dichromate dihydrate	0.602		9	-0.22	0.822		0.737		19	-0.13	0.258	
ECBC	0.603	14		-0.22		NA	0.784	14		-0.11		NA
FAL	0.657	37		-0.18		NA	0.851	36		-0.07		NA
IIVS	0.547	17		-0.26		NA	0.576	17		-0.24		NA
Sodium fluoride	79.8		22	1.90	0.016		47.4		15	1.68	0.313	
ECBC	61.3	9		1.79		0.007	48.7	14		1.69		NA
FAL	96.1	18		1.98		0.019	39.7	24		1.60		NA
IIVS	82.0	7		1.91		0.463	53.7	13		1.73		NA
Sodium hypochlorite	1211		57	3.08	0.040		1580		20	3.20	0.313	
ECBC	823	13		2.92		0.257	1863	31		3.27		NA
FAL	805	46		2.91		0.119	1243	46		3.09		NA
IIVS	2005	44		3.30		0.015	1633	11		3.21		NA
Sodium oxalate	40.8		23	1.61	0.643		355		1	2.55	0.926	
ECBC	42.0	41		1.62		NA	355	15		2.55		NA

Table 7-3 Reproducibility Results for the 3T3 and NHK NRU Test Methods

Reference Substance/Laboratory	3T3 NRU Test Method						NHK NRU Test Method					
	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]	Arithmetic Mean IC_{50} (µg/mL)[1]	Arithmetic IntraLab %CV	Arithmetic InterLab %CV	Log Arithmetic Mean IC_{50} (µg/mL)[1]	ANOVA P[2]	Contrast P[3]
FAL	31.0	28		1.49		NA	350	42		2.54		NA
IIVS	49.5	53		1.69		NA	360	26		2.56		NA
Sodium selenate	34.5		60	1.54	<0.001		11.2		40	1.05	0.134	
ECBC	12.7	13		1.10		<0.001	7.47	12		0.87		NA
FAL	54.2	19		1.73		< 0.001	16.1	59		1.21		NA
IIVS	36.5	14		1.56		0.026	10.0	13		1.00		NA
Strychnine	199		83	2.30	<0.001		69.3		39	1.84	0.364	
ECBC	389	21		2.59		<0.001	100	76		2.00		NA
FAL	124	16		2.09		0.018	52.5	53		1.72		NA
IIVS	83.5	6		1.92		<0.001	55.1	6		1.74		NA
Thallium Sulfate	7.50		72	0.88	0.165		0.16		23	-0.80	0.405	
ECBC	2.80	24		0.45		NA	0.198	51		-0.70		NA
FAL	13.4	78		1.13		NA	0.153	20		-0.82		NA
IIVS	6.30	28		0.80		NA	0.127	16		-0.90		NA
Trichloroacetic acid	928		27	2.97	0.005		427		24	2.63	0.134	
ECBC	762	13		2.88		0.022	348	18		2.54		NA
FAL	1220	6		3.09		0.002	541	28		2.73		NA
IIVS	801	14		2.90		0.069	394	13		2.60		NA
1,1,1-Trichloroethane	15538		52	4.19	<0.001		NA		NA	NA	NA	
ECBC	NA	NA		NA		NA	8137	7		3.91		NA
FAL	21250	11		4.33		NA	NA	NA		NA		NA
IIVS	9827	2		3.99		NA	NA	NA		NA		NA
Triethylenemelamine	0.568		135	-0.25	<0.001		1.95		12	0.29	0.562	
ECBC	0.086	11		-1.07		<0.001	1.69	57		0.23		NA
FAL	1.45	18		0.16		<0.001	2.03	23		0.31		NA
IIVS	0.169	29		-0.77		0.002	2.13	23		0.33		NA
Triphenyltin hydroxide	0.022		29	-1.66	0.688		0.013		55	-1.89	0.088	
ECBC	0.026	17		-1.59		NA	0.021	32		-1.68		NA
FAL	0.026	81		-1.59		NA	0.007	106		-2.15		NA
IIVS	0.015	55		-1.83		NA	0.011	32		-1.96		NA
Valproic acid	1177		76	3.07	< 0.001		533		28	2.73	0.081	
ECBC	547	12		2.74		NA	468	25		2.67		0.331
FAL	1807	10		3.26		NA	702	23		2.85		0.032
IIVS	NA	NA		NA		NA	430	17		2.63		0.135
Verapamil HCl	35.2		10	1.55	0.230		68.7		14	1.84	0.624	
ECBC	32.0	18		1.51		NA	60.5	22		1.78		NA
FAL	34.6	5		1.54		NA	79.4	42		1.90		NA
IIVS	38.9	11		1.59		NA	66.2	8		1.82		NA
Xylene	NA		NA	NA	NA	NA	NA		NA	NA	NA	
ECBC	NA	NA		NA		NA	NA	NA		NA		NA
FAL	NA	NA		NA		NA	NA	NA		NA		NA
IIVS	724	12		2.86		NA	486	38		2.69		NA

[1]Results reported on the same row with chemical names are the means of all the laboratories. Results reported on the same row as laboratories are the laboratory means.
[2]$p < 0.01$ indicated statistical significance.
[3]Contrasts were performed if ANOVA was significant ($p < 0.01$) to determine which laboratory was different from the other two laboratories. Significant contrasts were denoted by $p < 0.01$. If only two laboratories reported results, no contrast tests were necessary.
Abbreviations: Laboratories: ECBC- U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences. NA - no acceptable IC_{50} results reported or calculation was not performed (e.g., for contrast results).

158 **Table 7-4 Reference Substances with Significant Differences between Laboratories**
159 **for 3T3 NRU Test Method Results**

Reference Substance	Significant Contrast Results[1]			Insoluble/ Volatile[2]
	ECBC	FAL	IIVS	
Acetylsalicylic acid		H	L	
Arsenic trioxide		L	H	Precipitate
Busulfan		H		
Chloral hydrate	L	H		
Chloramphenicol	L	H		
Citric acid	L	H		
Colchicine	L	H		
Cupric sulfate pentahydrate	X	H	L	
Dibutyl phthalate		H	L	Precipitate
Dichlorvos	L	H		Precipitate
Disulfoton[3]				Precipitate
Ethylene glycol	L			
Glutethimide		H	L	
Haloperidol		H		
Meprobamate	L	H	X	
Phenylthiourea	L	H		
Potassium cyanide	L	H	X	Precipitate /Volatility
Procainamide HCl	L		H	
2-Propanol	L	X	H	Volatility
Propylparaben		H	L	
Sodium selenate	L	H		
Strychnine	H		L	Precipitate
Trichloroacetic acid		H		
1,1,1-Trichloroethane[4]				Precipitate
Triethylenemelamine	L	H		
Valproic acid[5]				Precipitate

160 [1]Laboratories significantly different from the other two at $p < 0.01$. H – Laboratory reported the highest mean
161 IC_{50}. L – Laboratory reported the lowest mean IC_{50}. X – Laboratory reported a mean IC_{50} between the values
162 of the other two laboratories.
163 [2]From **Table 5-8**. Precipitate reported by at least one laboratory is indicated by "Precipitate". Use of plate
164 sealers by at least one laboratory to prevent volatile contamination of control wells indicated by "Volatility".
165 [3]Significant ANOVA ($p < 0.01$), but no contrast analysis since only two laboratories (ECBC and IIVS) reported
166 IC_{50} values.
167 [4]Significant ANOVA ($p < 0.01$), but no contrast results since only two laboratories (FAL and IIVS) reported
168 IC_{50} values.
169 [5]Significant ANOVA ($p < 0.01$), but no contrast results since only two laboratories (ECBC and FAL) reported
170 IC_{50} values.
171 Laboratories: ECBC- U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives
172 Laboratory; IIVS – Institute for In Vitro Sciences.
173
174
175 For the 26 substances that yielded significantly different results among the laboratories,
176 contrast analyses indicated that ECBC and FAL were frequently different from the other
177 laboratories. ECBC tended to report the lowest IC_{50} values among the laboratories while

FAL tended to report the highest values of the three laboratories. ECBC reported significantly different results from the other two laboratories for 15 of the 26 substances. For 13 of the 15 substances, ECBC reported the lowest mean IC_{50} value among the three laboratories. FAL reported significantly different results from the other two laboratories for 20 of the 26 substances. For 18 of the 20 substances, FAL reported the highest mean IC_{50} value among the three laboratories. IIVS reported significantly different results for 11 of the 26 substances, with no great majority of highest or lowest IC_{50} values.

Differences Among the Laboratories for the NHK NRU Test Method

The ANOVA results in **Table 7-3** indicate that there were statistically significant ($p < 0.01$) laboratory differences for seven substances. These substances are listed in **Table 7-5** along with columns showing the laboratory statistically significantly differing from the other two laboratories (as indicated by the contrast results), and indications of whether any laboratory reported insolubility or volatility during conduct of the assay. Insolubility was reported for three of the seven substances.

Table 7-5 Reference Substances with Significant Differences between Laboratories for NHK NRU Test Method Results

Reference Substance	Significant Contrast Results[1]			Solubility/ Volatility[2]
	ECBC	FAL	IIVS	
Arsenic trioxide		L	H	Precipitate
Citric acid	H	L		Precipitate
Digoxin	H	L		
Dimethylformamide	H		L	
Disulfoton	L	H		Precipitate
Methanol[3]				
Propranolol HCl		H	L	

[1]Laboratories significantly different from the other two at $p < 0.01$. H – Laboratory reported the highest mean IC_{50}. L – Laboratory reported the lowest mean IC_{50}. X – Laboratory reported a mean IC_{50} between the values of the other two laboratories.
[2]From **Table 5-8**. Precipitate reported by at least one laboratory is indicated by "Precipitate". Use of plate sealers by at least one laboratory to prevent volatile contamination of control wells indicated by "Volatility".
[3]Significant ANOVA ($p < 0.01$), but no contrast results since only two laboratories (FAL and IIVS) reported IC_{50} values.
Laboratories: ECBC – U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.

For the seven substances that yielded significantly different results among the laboratories, ECBC and FAL were frequently different from the other laboratories. ECBC tended to report the highest IC_{50} value among the laboratories (4/7 substances) while FAL tended to report the lowest values among the three laboratories (3/7 substances).

7.2.2 CV Results for the 3T3 and NHK NRU Test Methods

CV was calculated as described in Section **5.3.3**. **Table 7-3** provides the intra- and inter-laboratory CV values for individual substances. **Table 7-6** summarizes the CV results for each test method. **Table 7-6** shows that median and mean CV values were often similar. Median CV values appeared always lower than the corresponding means, which indicated that large individual CV values skewed the CV distributions somewhat to the right.

Intralaboratory CV

Table 7-6 shows that both test methods had similar ranges for the intralaboratory CV. The mean intralaboratory CV values were the same, 26%. The median intralaboratory CVs were also similar: 23% for the 3T3 NRU test method and 24% for the NHK NRU test method. Of the three laboratories, FAL had the highest mean and median CV for both test methods and IIVS had the lowest mean and median CV for both test methods.

Interlaboratory CV

The mean and median interlaboratory CV for the reference substances was lower for the NHK NRU test method (mean = 28%; median = 21%) than for the 3T3 NRU test method (mean = 46%; median = 40%) (see **Table 7-6**).

Table 7-6 Summary of CV Results for the 3T3 and NHK NRU Test Methods

CV	3T3 NRU Test Method				NHK NRU Test Method			
	N	Mean	Median	Range	N	Mean	Median	Range
Intralaboratory CV	202	26%	23%	1-122%	208	26%	24%	1-129%
ECBC	68	23%	17%	2-95%	69	23%	19%	2-76%
FAL	66	33%	30%	1-98%	69	42%	32%	1-129%
IIVS	68	21%	13%	1-122%	70	14%	13%	1-38%
Interlaboratory CV	68	46%	40%	2-135%	68	28%	21%	1-99%

Abbreviations: N- number of values. Laboratories: ECBC- U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.
Note: For the 3T3 NRU test method, the following laboratories/substances did not obtain sufficient IC_{50} data for the calculation of an intralaboratory CV: carbon tetrachloride at any laboratory; disulfoton at FAL; gibberellic acid at FAL; lithium carbonate at FAL and IIVS; methanol at any laboratory; 1,1,1-trichloroethane at ECBC; valproic acid at IIVS; and xylene at ECBC and FAL. For the NHK assay, the following laboratories/substances did not obtain sufficient IC_{50} data for the calculation of an intralaboratory CV: carbon tetrachloride at any laboratory; methanol at ECBC; 1,1,1-trichloroethane at FAL and IIVS; and xylene at ECBC and FAL. For the 3T3 NRU test method, the following substances did not obtain sufficient IC_{50} data for the calculation of an interlaboratory CV: carbon tetrachloride, lithium carbonate; methanol; and xylene. For the NHK assay, the following substances did not yield sufficient IC_{50} data for the calculation of an interlaboratory CV: carbon tetrachloride; 1,1,1-trichloroethane; and xylene.

Variation of CV with Chemical Property

To identify the chemical characteristics that may yield high or low CV values, CV values were analyzed to determine their association with the following chemical attributes: physical state (i.e., solid or liquid), solubility, volatility, chemical class, GHS acute oral toxicity class (UN 2005), molecular weight, log Kow, IC_{50}, and boiling point. For categorical characteristics such as physical form, solubility (i.e., precipitate/no precipitate), volatile/not volatile, and chemical class, the mean CV values and CV ranges for the groups were compared to one another and to the overall mean CV and CV range for each test method. No statistical analyses were performed. For chemical characteristics measured by continuous variables, such as molecular weight, log K_{ow}, and IC_{50}, and boiling point, Spearman correlation analyses were performed.

Results of Intralaboratory CV Analysis

Table 7-7 shows the analysis of intralaboratory CV. The analysis of intralaboratory CV uses one mean intralaboratory CV for each reference substances that was calculated from the

intralaboratory CV values from each laboratory. With the exception of the amides, which had relatively low intralaboratory CV values (for both 3T3 and NHK NRU test methods), and organophosphates and halogenated hydrocarbons (for the 3T3 NRU test method only), which had relatively high intralaboratory CV values, there seemed to be little difference in CV values for the categorical physical/chemical/toxicological attributes. The mean intralaboratory CV values for solids and liquids were similar (26 vs. 24% for the 3T3 NRU test method; 27 vs. 23% for the NHK NRU test method). The mean intralaboratory CV values for reference substances for which precipitates were observed were similar to the mean intralaboratory CV values for substances for which no precipitates were observed (29 vs. 23% for the 3T3 NRU test method; 24 vs. 27% for the NHK NRU test method). The mean intralaboratory CV values for substances that exhibited volatility (i.e., indicated by laboratory use of film plate sealers to prevent contamination of control wells) were relatively similar to those that did not (31 vs. 24% for the 3T3 NRU test method; 27 vs. 26% for the NHK NRU test method). Similarly, the substances grouped by GHS toxicity category (UN 2005) had mean intralaboratory CV values that were similar (19-33% for the 3T3 NRU test method; 18-31% for the NHK NRU test method) to the overall mean CV values (26% for both the 3T3 and NHK NRU test methods).

Reference substances in the amide chemical class had unusually low mean intralaboratory CV values for both the 3T3 NRU test method (13%) and NHK NRU test method (10%) compared with the overall mean CV (26% for both test methods), but there were only three substances in the class (acetaminophen, dimethylformamide, and procainamide HCl). Reference substances in the organophosphate chemical class had unusually high mean intralaboratory CV values for the 3T3 NRU test method (46%), but not for the NHK NRU test method (26%) compared with the overall mean CV (26% for the 3T3 and NHK NRU test methods). There were only three substances in the class (dichlorvos, disulfoton, and parathion), but two of the three substances had relatively high mean intralaboratory CV values (17, 48 and 71%). Halogenated hydrocarbons also had high mean intralaboratory CV for the 3T3 NRU test method (46%), but not for the NHK NRU test method (14%) compared with the overall mean intralaboratory CV for each test method (26%). However, the mean intralaboratory CV for the 3T3 NRU test method was calculated from only two values; 7%

for 1,1,1-trichloroethane and 84% for lindane. No laboratory obtained sufficient toxicity for the calculation of an IC_{50} for the carbon tetrachloride, the third halogenated hydrocarbon.

Table 7-7 Intralaboratory CV by Chemical Characteristics for the 3T3 and NHK NRU Test Methods

Class/Attribute	3T3 NRU Test Method			NHK NRU Test Method		
	N[a]	Range	Mean	N[b]	Range	Mean
All chemicals	70	1-122%	26%	71	1-129%	26%
Chemical form						
Solid	53	4-84	26	53	6-50	27
Liquid	17	6-71	24	18	2-40	23
Solubility						
Precipitate[c]	24	7-84	29	2[a]	2-47	24
No precipitate	46	4-55	23	50	7-57	27
Volatility[d]						
Volatile	10	6-84	31	9	11-50	27
Nonvolatile	62	4-71	24	63[b]	2-57	26
Chemical Class						
Alcohols	9	6-42	22	10	10-37	21
Carboxylic acids	12	10-41	20	12	7-48	26
Heterocyclics	14	6-59	30	14	13-50	31
Organophosphorous	3	17-71	46	3	20-32	26
Amides	3	4-28	13	3	2-16	10
Halogenated hydrocarbons	2	7-84	46	2	7-21	14
Inorganics	15	9-43	24	15	6-50	29
Toxicity Class						
≤ 5 mg/kg	7	9-71	33	7	20-40	30
> 5 - ≤ 50	12	13-59	32	12	12-50	31
> 50 - ≤ 300	12	11-84	33	12	17-37	25
> 300 - ≤ 2000	16	4-51	21	16	6-57	25
> 2000 - ≤ 5000	10[a]	9-32	19	10[a]	7-50	31
> 5000	13[b]	6-42	19	14	2-40	18
Correlations	**N**	**r_s**	**P value**	**N**	**r_s**	**P value**
Molecular weight	70[a,b]	0.323	0.006	71[b]	0.199	0.097
Log K_{ow}	50[e]	0.117	0.421	51[e]	0.311	0.026
IC_{50}	70[a,b]	-0.436	0.0002	71[b]	-0.362	0.002
Boiling point	27	0.576	0.002	28	0.277	0.154

[a]One intralaboratory CV for each chemical was calculated by averaging the CV values for the laboratories that reported sufficient data for the calculation of a CV. No CV was calculable for carbon tetrachloride or methanol.
[b]One intralaboratory CV for each chemical was calculated by averaging the CV values for the laboratories that reported sufficient data for the calculation of a CV. No CV was calculable for carbon tetrachloride.
[c]Denoted by laboratory reports of precipitate in the stock reference substance solutions or in cell culture (see **Table 5-8**).
[d]Denoted by laboratory reports of using plate sealers to avoid contamination of the VC wells (see **Table 5-8**).
[e]Number of reference substances with CV values and log K_{ow} data.
[f]Number of reference substances with CV values and boiling point data.

For the characteristics amenable to correlation analysis, none of the correlation coefficients were large (absolute value of $r_s < 0.6$), but several were statistically significantly different from zero for the 3T3 NRU test method. Molecular weight ($p = 0.006$), IC_{50} ($p = 0.0002$), and boiling point ($p = 0.002$) exhibited statistically significant correlations ($p < 0.05$) to intralaboratory CV for the 3T3 NRU test method. For molecular weight, the higher molecular weight substances had higher intralaboratory CV values. For IC_{50}, however, the substances with lower IC_{50} values had higher CV values. The inverse correlation between intralaboratory CV values and IC_{50} is consistent with the common observation that measurements with very low values tend to have high CV values. The fact that substances with higher boiling points had higher CV values was consistent with the categorical analysis of volatility. The substances that exhibited volatile characteristics (i.e., high reference substance concentration wells contaminated the VC wells) in the 3T3 NRU test method had higher mean intralaboratory CV values (31%) than the substances that did not exhibit volatile characteristics (24%), but the difference did not seem large.

Likewise, for the NHK NRU test method, two of the characteristics amenable to correlation analysis were statistically significantly different from zero, but the correlation coefficients did not have large magnitudes (absolute value of $r_s < 0.4$). Log K_{ow} ($p = 0.026$) and IC_{50} ($p = 0.002$) exhibited statistically significant correlations ($p < 0.05$) to intralaboratory CV for the NHK NRU test method. Log K_{ow} was positively correlated to the mean intralaboratory CV for each substance, but IC_{50}, was negatively correlated to the mean IC_{50} for each substance.

Results of Interlaboratory CV Analysis

Table 7-8 shows the analysis of interlaboratory CV. With the exception of chemical class, there seemed to be little difference in interlaboratory CV values for most of the categorical physical/chemical characteristics. The mean interlaboratory CV values for solids and liquids were similar (48 vs. 46% for the 3T3 NRU test method and 28 vs. 27% for the NHK NRU test method). The mean interlaboratory CV values for substances for which precipitates were observed was similar to the mean interlaboratory CV values for substances for which no

precipitates were observed (56 vs. 43% for the 3T3 NRU test method and 29 vs. 28% for the NHK NRU test method). The mean interlaboratory CV values for substances that exhibited volatile characteristics appeared similar to those that did not (51 vs. 46% for the 3T3 NRU test method and 32 vs. 28% for the NHK NRU test method).

Reference substances in the amide chemical class had unusually low mean interlaboratory CV values for both the 3T3 NRU test method (15%) and NHK NRU test method (16%) compared with the overall mean interlaboratory CV (46% for the 3T3 NRU test method and 28% for the NHK NRU test method). Chemicals in the organophosphate chemical class had unusually high mean interlaboratory CV values for the 3T3 NRU test method (74%) and moderately higher mean interlaboratory CV values for the NHK NRU test method (42%) compared with the overall mean interlaboratory CV (46% for the 3T3 NRU test method and 28% for the NHK NRU test method). The high mean interlaboratory CV value for organophosphates in the NHK NRU test method, however, was produced largely by the high interlaboratory CV of 99% for disulfoton. The interlaboratory CV values for dichlorvos and parathion were 20% and 8%, respectively. Heterocyclic compounds also had higher mean interlaboratory CV values for the 3T3 NRU test method but not for the NHK NRU test method. As a group, the 14 heterocyclic compounds had a mean interlaboratory CV of 61% while the overall mean interlaboratory CV for the 3T3 NRU test method was 46%. Although there were a few low CV values (e.g., 8, 18) in the heterocyclic group, there were seven values greater than the overall mean CV of 46%. The median interlaboratory CV for the heterocyclic group was 52%.

Table 7-8 Interlaboratory CV by Chemical Characteristics for the 3T3 and NHK NRU Test Methods

Class/Attribute	3T3 NRU Test Method			NHK NRU Test Method		
	N	Range	Mean	N	Range	Mean
All chemicals	68[a]	2-135%	46%	69[b]	1-99%	28%
Chemical Form						
Solids	52	3-135	48	53	1-91	28
Liquids	16	6-124	46	16	1-99	27
Solubility						
Precipitate[c]	22	3-127	56	19	1-99	29
No precipitate	47	3-135	43	50	1-88	28
Volatility						
Volatile[d]	10	21-127	51	9	8-86	32
Nonvolatile	58	3-135	46	60	1-99	28
Chemical Class						
Alcohols	9	12-119	38	10	11-42	22
Carboxylic acids	12	12-124	46	12	1-61	27
Heterocyclics	14	8-135	61	14	5-85	32
Organophosphorous	3	57-111	74	3	8-99	42
Amides	3	6-28	15	3	13-19	16
Halogenated hydrocarbons	2	52-58	55	1	20	20
Inorganics	14	3-127	48	15	4-91	29
Toxicity Class						
≤ 5 mg/kg	7	12-135	69	7	12-99	37
> 5 - ≤ 50	12	33-127	78	12	8-91	41
> 50 - ≤ 300	12	8-120	37	12	10-41	26
> 300 - ≤ 2000	15	11-85	38	15	1-61	20
> 2000 - ≤ 5000	9	3-69	29	9	1-85	27
> 5000	13	3-124	39	13	2-44	25
Correlations		**r_s**	**P value**		**r_s**	**P value**
Molecular weight	68	0.193	0.115	69	0.136	0.265
Log K_{ow}	49[e]	0.194	0.182	49	0.170	0.244
IC_{50}	68	-0.295	0.015	69	-0.271	0.024
Boiling point	24[f]	0.467	0.021	26	-0.131	0.525

[a]The following chemicals did not have sufficient IC_{50} data for the calculation of an interlaboratory CV: carbon tetrachloride, lithium carbonate; methanol; and xylene.

[b]The following substances did not yield sufficient IC_{50} data for the calculation of an interlaboratory CV: carbon tetrachloride; 1,1,1-trichloroethane; and xylene.

[c]Denoted by laboratory reports of precipitate in the stock reference substance solutions or in cell culture (see **Table 5-8**).

[d]Denoted by laboratory reports of using plate sealers to avoid contamination of the VC wells (see **Table 5-8**).

[e]Number of reference substances with CV values and log K_{ow} data.

[f]Number of reference substances with CV values and boiling point data.

Mean interlaboratory CV values tended to be large for chemicals in the most toxic GHS acute categories, especially for the 3T3 NRU test method. For the 3T3 NRU test method, the mean interlaboratory CV for chemicals in the classes for $LD_{50} \leq 5$ mg/kg (69%) and $5 < LD_{50}$

≤ 50 mg/kg (78%) were much larger than the mean overall interlaboratory CV (46%,). For the NHK NRU test method, the mean interlaboratory CV for chemicals in the classes for 5 < LD_{50} ≤ 5 mg/kg (37%) and 5 < LD_{50} ≤ 50 mg/kg (41%) were much larger than the mean overall interlaboratory CV (28%).

For the characteristics amenable to correlation analysis, none of the correlation coefficients were large (absolute value of r_s < 0.5), but IC_{50} ($p = 0.015$) and boiling point ($p = 0.021$) exhibited statistically significant correlations ($p < 0.05$) to interlaboratory CV for the 3T3 NRU test method. There was a negative correlation between interlaboratory CV and IC_{50}, but the correlation between boiling point and interlaboratory CV was positive. The positive correlation of CV with boiling point was largely consistent with the categorical analysis of volatility. The substances that exhibited volatile characteristics in the 3T3 NRU test method had slightly higher mean CV than for the substances that did not exhibit volatile characteristics (51 vs. 46%). For the NHK NRU test method, only IC_{50} was significantly correlated ($p = 0.024$) to interlaboratory CV with a negative correlation ($r_s = -0.271$).

7.2.3 Comparison of Laboratory-Specific Linear Regression Analyses for the Prediction of *In Vivo* Rodent LD_{50} Values from *In Vitro* NRU IC_{50} Values

The laboratory-specific regressions presented in **Table 6-1** of **Section 6.1.1** were compared to one another (for each test method) with a goodness of fit F-test as described in **Section 5.3.3.** The comparisons indicated that the laboratory-specific regressions for both test methods were not significantly different ($p < 0.05$) from one another. The comparison of the laboratory-specific 3T3 NRU regressions to one another yielded $p = 0.796$. The comparison of the laboratory-specific NHK NRU regressions to one another yielded $p = 0.985$. Because the laboratory-specific regressions were not statistically different, data were combined into a single regression for each test method using a geometric mean of the laboratory-specific IC_{50} values for each substance (see **Section 6.1.1**).

7.2.4 Laboratory Concordance for the Prediction of GHS Acute Oral Toxicity Category

This section provides the percentage of substances for which the laboratory-specific IC_{50} data yielded the same (for all three laboratories) GHS toxicity categorization when used with the

regressions evaluated in **Sections 6.3.1** through **6.3.3**. Data for the same reference substances for each test method were evaluated to determine the laboratory concordance for each regression. Forty-three substances were evaluated for the 3T3 NRU test method and 44 substances were evaluated for the NHK NRU test method. Of the original 72 substances tested, epinephrine bitartrate, colchicine, and propylparaben were excluded from all analyses because they were removed from the calculation of the RC rat-only weight regressions due to the lack of rat oral reference LD_{50} data. The 21 substances with specific mechanisms of toxicity in **Table 6-3** were excluded from all analyses to be consistent with those removed from the RC rat-only weight regression excluding substances with specific mechanisms of toxicity. These substances have known mechanisms of toxicity that are not expected to be active in the 3T3 or NHK cell cultures. Carbon tetrachloride, methanol, gibberellic acid, lithium carbonate, and xylene were excluded from the 3T3 NRU evaluations because at least one laboratory failed to attain sufficient toxicity in any test for the calculation of an IC_{50}. Carbon tetrachloride, methanol, 1,1,1-trichloroethane, and xylene were excluded from the NHK NRU analyses because at least one laboratory failed to attain sufficient toxicity in any test for the calculation of an IC_{50}.

Laboratory Concordance for the 3T3 and NHK NRU Test Methods with the RC Millimole Regression

Appendix J (**Table J-1** for the 3T3 NRU test method and **Table J-3** for the NHK NRU test method) shows the laboratory concordance of the observed (i.e., *in vivo* categories for the initial LD_{50} values in **Table 3-2**) and predicted GHS toxicity categories for each substance determined in each *in vitro* NRU cytotoxicity test method using the laboratory-specific geometric mean IC_{50} values and the RC millimole regression, log LD_{50} (mmol/kg) = 0.435 x log IC_{50} (mM) + 0.625. The observed LD_{50} values are the rodent LD_{50} values from **Table 3-2.**

For the 43 substances that yielded IC_{50} results in all laboratories using the 3T3 NRU test method, the laboratories agreed on the GHS toxicity category for 31 substances (72%). The 12 substances that produced discordant results among the laboratories were cupric sulfate pentahydrate, cycloheximide, dimethylformamide, diquat dibromide, phenol, phenylthiourea,

sodium arsenite, sodium oxalate, sodium selenate, thallium sulfate, triethylenemelamine, and 1,1,1-trichloroethane. The laboratory predictions for these substances disagreed by one GHS toxicity category.

For the 44 substances that yielded IC_{50} results in all laboratories using the NHK NRU test method, the laboratories agreed on toxicity category for 39 substances (89%). The five substances that produced discordant results among the laboratories were arsenic trioxide, digoxin, ethanol, 2-propanol, and sodium arsenite. The laboratory predictions for these substances disagreed by one toxicity category. Laboratory concordance was greater for the NHK assay than for the 3T3 assay (89% vs 72%).

Laboratory Concordance of the 3T3 and NHK NRU Test Methods with the RC Rat-Only Weight Regression

Appendix J (**Table J-5** for the 3T3 NRU test method and **Table J-6** for the NHK NRU test method) shows the laboratory concordance of the observed (i.e., *in vivo* reference categories for LD_{50} values in **Table 4-2**) and predicted GHS toxicity categories for each substance as determined for each test method using the laboratory-specific geometric mean IC_{50} in the RC rat-only weight regression, log LD_{50} (mg/kg) = log IC_{50} (μg/mL) x 0.372 + 2.024, from **Table 6-2.**

For the 43 substances that yielded IC_{50} results in all laboratories using the 3T3 NRU test method, the laboratories agreed on the GHS toxicity category for 34 substances (79%). The nine substances that produced discordant results among the laboratories were boric acid, cupric sulfate pentahydrate, cycloheximide, 2-propanol, propranolol HCl, sodium selenate, thallium sulfate, triethylenemelamine, and 1,1,1-trichloroethane. The laboratory predictions for these substances disagreed by one GHS toxicity category.

For the 44 substances that yielded IC_{50} results in all laboratories using the NHK NRU test method, the laboratories agreed on toxicity category for 39 substances (89%). The five substances that produced discordant results among the laboratories were arsenic trioxide, digoxin, glycerol, sodium chloride, and thallium sulfate. The laboratory predictions for these

substances disagreed by one toxicity category. Laboratory concordance was greater for the NHK assay than for the 3T3 assay (89% vs 79%).

Laboratory Concordance of the 3T3 and NHK NRU Test Methods with the RC Rat-Only Weight Regression Excluding Substances with Specific Mechanisms of Toxicity

Appendix J (**Table J-7** for the 3T3 NRU test method and **Table J-8** for the NHK NRU test method) shows the laboratory concordance of the observed (i.e., *in vivo*) and predicted GHS toxicity categories for each substance as determined for each test method using the laboratory-specific geometric mean IC_{50} values in the RC rat-only weight regression after exclusion of substances with specific mechanisms of toxicity, log LD_{50} (mg/kg) = log IC_{50} (μg/mL) x 0.357 + 2.194 (**Table 6-2**).

For the 43 substances considered in the analysis of the 3T3 NRU test method, the three laboratories agreed on the toxicity category for 36 (84%) of the substances. The seven substances that produced discordant results among the laboratories were boric acid, cupric sulfate pentahydrate, diquat dibromide, sodium hypochlorite, thallium sulfate, 1,1,1-trichloroethane, and valproic acid. The laboratory predictions for these substances disagreed by one GHS toxicity category.

The extent of laboratory concordance for the RC rat-only weight regression after excluding substances with specific mechanisms of toxicity was the same for the NHK NRU test method (i.e., 84%, 37/44). The seven substances that produced discordant results among the laboratories were arsenic trioxide, digoxin, glycerol, hexachlorophene, mercury chloride, sodium chloride, and sodium hypochlorite. The laboratory predictions for these substances disagreed by one GHS toxicity category.

7.3 Historical Positive Control Data

The reproducibility of the positive control (SLS) data was assessed by CV analysis, ANOVA, and linear regression over time as described in **Section 5.3.4**. The SLS data analyzed for variability are slightly different from those used to determine the PC acceptance

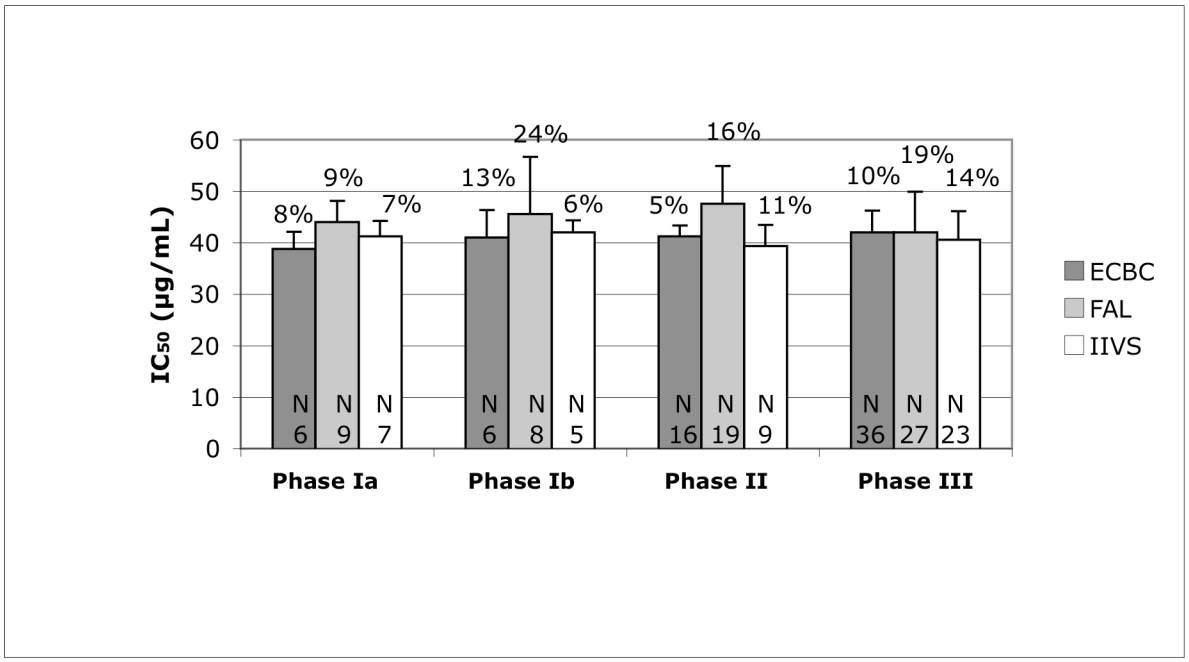
limits shown in **Table 5-2**. To get an assessment of the true variation of SLS IC_{50} values, the reproducibility analyses included IC_{50} values from SLS tests that failed the test acceptance criterion for the IC_{50} acceptance limits determined for each study phase. These additional SLS tests, however, passed all other test acceptance criteria. If more than one SLS test was performed in a single day (for each test method and laboratory), the IC_{50} values were averaged to determine a single IC_{50} for the day so that multiple results from a single day would not overly influence the average for each phase.

Figure 7-1 shows the average SLS IC_{50} values for each test method, laboratory, and study phase. Graphically, it appears that the SLS IC_{50} for the 3T3 NRU test method was relatively consistent over the entire period of the study (approximately 2.5 years). The intralaboratory CV values (shown in **Figure 7-1**) for the individual study phases ranged from 5% to 24%. With the exception of the Phase Ib CV at FAL, the CV values for each laboratory and phase were less than 20%. The interlaboratory CV values were even smaller: 6% for Phases Ia and Ib; 10% for Phase II; and 2% for Phase III.

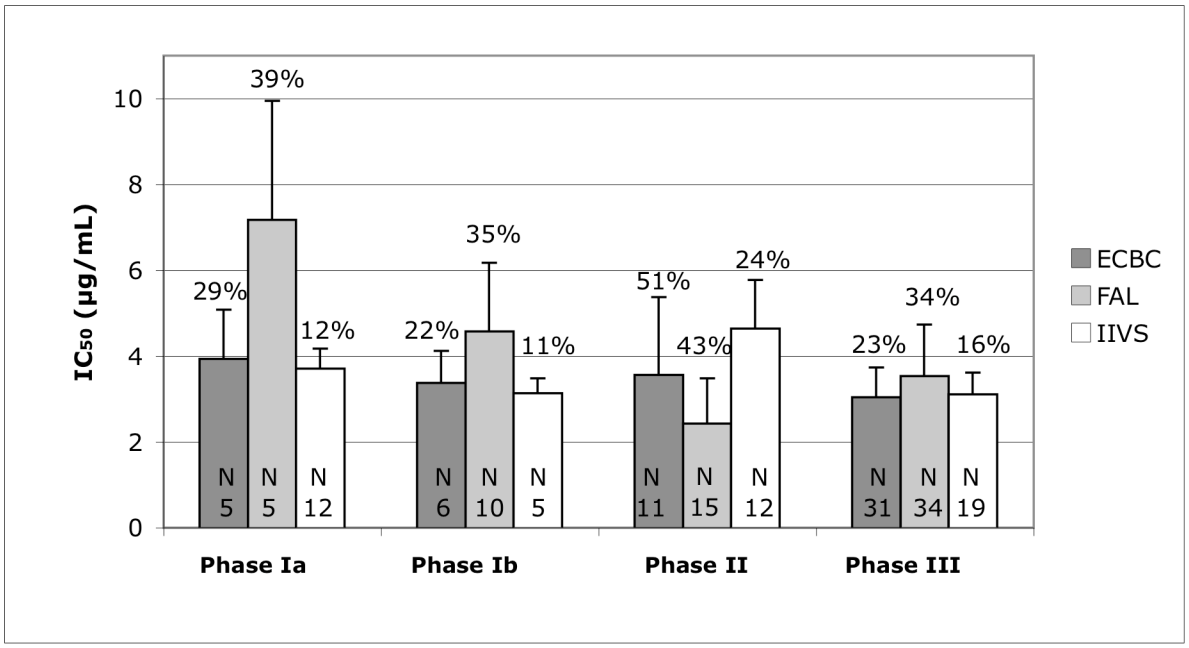
Figure 7-1 shows that the SLS IC_{50} for the NHK NRU test method tended to vary with time, but, with the exception of the SLS IC_{50} results from FAL, there appeared to be no consistent trend. The IC_{50} values from FAL, which changed NHK cell culture methods after Phase Ib (see **Section 5.1.3**), tended to decrease over time. Although the change in cell culture methods reduced the magnitude of the IC_{50}, the variability (as evidenced by the intralaboratory CV values shown in **Figure 7-1**) remained relatively high (CV $\geq$ 34% for all FAL study phases). The CV values for all the laboratories and study phases indicated that the SLS IC_{50} values for the NHK NRU test method was more variable within laboratories than the SLS IC_{50} for 3T3 NRU test method. CV values for the SLS IC_{50} for the NHK NRU test method ranged from 11 to 51%, with nine of the 12 values greater than 20%. The interlaboratory CV values, which were also greater than those for the 3T3 NRU test method, were: 39% for Phase Ia; 21% for Phase Ib; and 31% for Phase II; and 8% for Phase III.

Figure 7-1 SLS IC_{50} for Each Laboratory and Study Phase

a 3T3 NRU Test Method



b NHK NRU Test Method



Bars show mean IC_{50} values. Error bars show standard deviation. Percent values above error bars are intralaboratory CVs.
Laboratories: ECBC- U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.

7.3.1 ANOVA and Linear Regression Results for the 3T3 NRU Test Method

SLS IC_{50} Variation with Time

Table 7-9 shows the ANOVA results for SLS from the 3T3 NRU test method. When the IC_{50} values within each laboratory were compared by study phase (i.e., the ANOVA factor was study phase), there were no statistically significant differences ($p < 0.01$) between study phases for any laboratory. **Table 7-10** shows that the slopes of the linear regressions of the IC_{50} values over time (expressed as index values) were statistically different from zero for ECBC and FAL ($p = 0.001$ and 0.012, respectively). Since the slopes were so small (0.000204 and -0.000324), they were considered to be unimportant. The slope of the IIVS regression of SLS IC_{50} over time was not statistically different from zero ($p = 0.651$; **Table 7-10**), which was entirely consistent with the ANOVA (**Table 7-9**) indicating that SLS IC_{50} from IIVS did not vary with study phase ($p = 0.854$). The ANOVA with study phase as the factor (with laboratories combined) indicated that the 3T3 NRU IC_{50} values from all the laboratories were consistent over time since data from the various study phases were not statistically significantly different ($p = 0.304$).

Comparison of SLS IC_{50} Among the Laboratories

When all study phases from each laboratory were combined, ANOVA, with laboratory as the factor, indicated that the SLS IC_{50} for the 3T3 NRU test method differed in some statistically significant fashion among the laboratories ($p < 0.006$). However, the differences between laboratories look rather small in **Figure 7-1** since the SDs for the laboratories clearly overlap one another.

560 **Table 7-9 ANOVA Results for SLS IC_{50} from the 3T3 NRU Test Method**

Study Phase/ Laboratory	ECBC				FAL				IIVS			
	Log Mean IC_{50} (mM)	SD	N	P[1]	Log Mean IC_{50} (mM)	SD	N	P[1]	Log Mean IC_{50} (mM)	SD	N	P[1]
Test for differences between phases within each laboratory												
Phase Ia	-0.876	0.042	6	0.031	-0.811	0.046	9	0.015	-0.850	0.034	7	0.854
Phase Ib	-0.864	0.066	6		-0.846	0.065	8		-0.838	0.025	5	
Phase II	-0.848	0.027	16		-0.796	0.057	19		-0.854	0.025	8	
Phase III	-0.842	0.036	36		-0.851	0.066	27		-0.844	0.041	23	
Test for differences between laboratories (phases combined)												
All Phases	-0.849	0.039	64	0.006	-0.826	0.062	63		-0.847	0.035	44	
Test for differences between phases (laboratories combined)												
Phase Ia	-0.839	0.049	22	0.304								
Phase Ib	-0.850	0.056	19									
Phase II	-0.831	0.047	34									
Phase III	0.845	0.045	86									

561 [1]Statistically significant at $p < 0.01$.
562 Abbreviations: N- number of values; SD – standard deviation. Laboratories: ECBC- U.S. Army Edgewood
563 Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.
564

Table 7-10 Linear Regression Analysis of SLS IC_{50} Over Time[1]

Test Method/ Laboratory	Slope	P-value (Slope)[2]	Intercept
3T3 NRU			
ECBC	0.000204	0.001	-0.874
FAL	-0.000324	0.012	-0.796
IIVS	0.0000304	0.651	-0.850
NHK NRU			
ECBC	-0.000559	0.002	-1.901
FAL	-0.00112	< 0.001	-1.737
IIVS	-0.000445	0.002	-1.885

[1]Time was expressed as index values. The index value of each test reflected the order of testing without respect to the time lapsing between tests.
[2]Statistically significant from zero at $p < 0.05$.
Laboratories: ECBC- U.S. Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.

7.3.2 ANOVA and Linear Regression Results for the NHK NRU Test Method

SLS IC_{50} Variation with Time

Table 7-11 shows the ANOVA results for the NHK NRU test method. When the IC_{50} values within each laboratory were compared by study phase (i.e., the ANOVA factor was phase), the phases were statistically different ($p < 0.01$) at each laboratory. The IC_{50} values from the various study phases were also significantly different from one another when the laboratory data were combined ($p < 0.001$). Linear regression analyses showed that the slopes for IC_{50} over time (expressed as an index values) were statistically significantly greater than zero for each laboratory (see **Table 7-10**). Since the slopes were so small (-0.000559, -0.00112, and -0.000445), they were considered to be unimportant.

583 **Table 7-11 ANOVA Results for SLS IC_{50} from the NHK NRU Test Method**

Study Phase/ Laboratory	ECBC				FAL				IIVS			
	Log Mean IC_{50} (mM)	SD	N	P[1]	Log Mean IC_{50} (mM)	SD	N	P[1]	Log Mean IC_{50} (mM)	SD	N	P[1]
Test for differences between phases within each laboratory												
Phase Ia	-1.867	0.135	5	0.001	-1.656	0.125	5	< 0.001	-1.904	0.060	12	< 0.001
Phase Ib	-1.936	0.092	6		-1.829	0.141	10		-1.965	0.046	5	
Phase II	-2.007	0.109	11		-1.982	0.173	15		-1.863	0.058	12	
Phase III	-1.990	0.098	31		-1.941	0.113	34		-1.972	0.070	19	
Test for differences between laboratories (phases combined)												
All Phases	-1.971	0.113	53	< 0.001	-1.879	0.175	64		-1.924	0.073	48	
Test for differences between phases (laboratories combined)												
Phase Ia	-1.833	0.143	22	< 0.001								
Phase Ib	-1.891	0.125	21									
Phase II	-1.964	0.139	38									
Phase III	-1.971	0.100	84									

584 [1]Statistically significant at $p < 0.01$.
585 Abbreviations: N- number of values; SD – standard deviation. Laboratories: ECBC – U.S. Army Edgewood
586 Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences.
587

Comparison of SLS IC_{50} Among the Laboratories

The ANOVA results, with laboratory as a factor (**Table 7-11)** indicated that the SLS IC_{50} was statistically different among the laboratories when the data from the study phases were pooled ($p < 0.001$). **Figure 7-1** shows that the SLS data from ECBC and IIVS were rather similar for Phases Ia, Ib, and III. The SLS IC_{50} data from FAL looks different from the other two laboratories for Phases Ia, Ib, and II, but the bars and SDs for Phase III show that the data from all laboratories were similar.

7.4 Laboratory Concordance for Solvent Selection

The solvents used to dissolve the reference substances are shown in **Table 7-12**. For Phases Ib and II, the SMT selected the solvents to use for cytotoxicity testing based on the solubility results provided by BioReliance (see **Table 5-7**) using the solubility protocol in **Appendix G2**. Despite the fact that the solubility of an individual substance in 3T3 medium and NHK medium might be different, the SMT chose the same solvent for both test methods, rather than choosing one for the 3T3 assay and one for the NHK assay. For example, if solubility in the 3T3 medium was ≥ 2 mg/mL and solubility in the NHK medium was < 2 mg/mL, and the substance was soluble in DMSO at 200 mg/mL, then the SMT selected DMSO as the solvent for cytotoxicity testing.

During Phases Ib and II, the SMT noted that BioReliance sometimes achieved higher solubility than the cytotoxicity laboratories (e.g., see the results for arsenic trioxide, aminopterin, and chloramphenicol in **Table 5-7**). In an attempt to avoid the selection of a solvent for which one or more laboratories could not achieve the desired solubility, the SMT used the solubility data from all the laboratories to determine solvent selections for cytotoxicity testing in Phase III. The SMT viewed BioReliance's NHK and 3T3 media solubility results for each substance in Phases Ib and II to be one result for media and took a similar approach in Phase III when considering all the laboratory results to determine the solvent to use for cytotoxicity testing. For example, if one laboratory had achieved solubility at 2 mg/mL in medium, but the other laboratories had not, and the substance was soluble in DMSO at 200 mg/mL, then the SMT selected DMSO as the solvent. **Table 7-12** shows that

cell culture medium was used to test as the solvent for 38 substances and DMSO was used as the solvent for 34 substances.

The solubility protocol used by the cytotoxicity laboratories failed to guide the selection of a solvent for five substances because they were insoluble at all concentrations tested in at least one laboratory. Arsenic trioxide was insoluble at all the cytotoxicity laboratories. IIVS also found sodium oxalate, strychnine, and triethylenemelamine insoluble in any solvent, and FAL found thallium sulfate insoluble in any solvent. To select a solvent for cytotoxicity testing of these substances, the SMT used the solubility results from the laboratories that did achieve solubility.

Table 7-12 Solvent Determinations by Laboratory

Reference Substance	Solvent for Testing[1]	ECBC	FAL	IIVS
Acetaminophen	DMSO	Medium	Medium	DMSO
Acetonitrile	Medium	Medium	Medium	Medium
Acetylsalicylic acid	DMSO	Medium	DMSO	Medium
Aminopterin	DMSO	DMSO	DMSO	DMSO
5-Aminosalicylic acid	Medium	Medium	Medium	Medium
Amitriptyline HCl	DMSO	DMSO	DMSO	DMSO
Arsenic III trioxide	Medium	ID	ID	ID
Atropine sulfate	Medium	Medium	Medium	Medium
Boric aid	Medium	Medium	Medium	Medium
Busulfan	DMSO	DMSO	DMSO	DMSO
Cadmium II chloride	DMSO	DMSO	DMSO	DMSO
Caffeine	Medium	Medium	Medium	Medium
Carbamazepine	DMSO	Medium	DMSO	DMSO
Carbon tetrachloride	DMSO	Medium	DMSO	Medium
Chloral hydrate	Medium	Medium	Medium	Medium
Chloramphenicol	DMSO	DMSO	DMSO	Medium
Citric acid	Medium	Medium	Medium	Medium
Colchicine	Medium	Medium	Medium	Medium
Cupric sulfate pentahydrate	Medium	Medium	Medium	Medium
Cycloheximide	Medium	Medium	Medium	Medium
Dibutyl phthalate	DMSO	DMSO	DMSO	DMSO
Dichlorvos (DDVP)	DMSO	Medium	DMSO	Medium
Diethyl phthalate	DMSO	DMSO	DMSO	DMSO
Digoxin	DMSO	DMSO	DMSO	DMSO
Dimethylformamide	Medium	Medium	Medium	Medium
Diquat dibromide monohydrate	Medium	Medium	Medium	Medium
Disulfoton	DMSO	DMSO	DMSO	DMSO
Endosulfan	DMSO	DMSO	DMSO	DMSO
Epinephrine bitartrate	Medium	Medium	Medium	Medium
Ethanol	Medium	Medium	Medium	Medium
Ethylene glycol	Medium	Medium	Medium	Medium
Fenpropathrin	DMSO	DMSO	DMSO	DMSO

Table 7-12 Solvent Determinations by Laboratory

Reference Substance	Solvent for Testing[1]	ECBC	FAL	IIVS
Gibberellic acid	Medium	Medium	Medium	Medium
Glutethimide	DMSO	DMSO	DMSO	DMSO
Glycerol	Medium	Medium	Medium	Medium
Haloperidol	DMSO	DMSO	DMSO	DMSO
Hexachlorophene	DMSO	DMSO	DMSO	DMSO
Lactic acid	Medium	Medium	Medium	Medium
Lindane	DMSO	DMSO	DMSO	DMSO
Lithium I carbonate	Medium	Medium	Medium	Medium
Meprobamate	DMSO	Medium	Medium	DMSO
Mercury II chloride	DMSO	DMSO	DMSO	DMSO
Methanol	DMSO	Medium	Medium	DMSO
Nicotine	Medium	Medium	Medium	Medium
Paraquat	Medium	Medium	Medium	Medium
Parathion	DMSO	DMSO	DMSO	DMSO
Phenobarbital	DMSO	Medium	DMSO	DMSO
Phenol	Medium	Medium	Medium	Medium
Phenylthiourea	DMSO	DMSO	Medium	DMSO
Physostigmine	DMSO	Medium	DMSO	DMSO
Potassium I chloride	Medium	Medium	Medium	Medium
Potassium cyanide	Medium	Medium	Medium	Medium
Procainamide HCl	Medium	Medium	Medium	Medium
2-Propanol	Medium	Medium	Medium	Medium
Propranolol HCl	DMSO	Medium	Medium	Medium
Propylparaben	DMSO	DMSO	DMSO	DMSO
Sodium arsenite	Medium	Medium	Medium	Medium
Sodium chloride	Medium	Medium	Medium	Medium
Sodium dichromate dihydrate	Medium	Medium	Medium	Medium
Sodium fluoride	Medium	Medium	Medium	Medium
Sodium hypochlorite	Medium	Medium	Medium	Medium
Sodium oxalate	Medium	Medium	Medium	ID
Sodium selenate	Medium	Medium	Medium	Medium
Strychnine	Medium	Medium	Medium	ID
Thallium I sulfate	Medium	Medium	ID	Medium
Trichloroacetic acid	Medium	Medium	Medium	Medium
1,1,1-Trichloroethane	Medium	Medium	Medium	Medium
Triethylenemelamine	DMSO	Medium	DMSO	ID
Triphenyltin hydroxide	DMSO	DMSO	DMSO	DMSO
Valproic acid	DMSO	Medium	DMSO	DMSO
Verapamil HCl	DMSO	DMSO	DMSO	DMSO
Xylene	DMSO	DMSO	DMSO	DMSO
DMSO Total	34	22	29	28
Medium Total	38	49	41	40

ID-insufficient data to select solvent.

[1]Solvents for testing as determined by the SMT and used in the study by each laboratory: Medium = cell culture medium; DMSO = dimethyl sulfoxide

ECBC – US Army Edgewood Chemical Biological Center; FAL – FRAME Alternatives Laboratory; IIVS – Institute for In Vitro Sciences

The cytotoxicity laboratories selected the same solvent for 55 of the 72 reference substances (76%). Excluding the five substances that were found to be insoluble in any solvent by at least one laboratory, there were 12 substances for which the cytotoxicity laboratories disagreed: acetaminophen, acetylsalicylic acid, carbamazepine, carbon tetrachloride, chloramphenicol, dichlorvos, meprobamate, methanol, phenobarbital, phenylthiourea, physostigmine, and valproic acid. Every laboratory reported relatively low solubility, ≤ 2 mg/mL, in medium for these substances. Since 2 mg/mL in medium is the departure point for the selection medium or DMSO, a small variation in results causes the laboratories to select different solvents. The solubility of acetaminophen, for example was reported as 2 mg/mL in culture media by ECBC and FAL, but < 2 mg/mL by IIVS. IIVS found it soluble in 200 mg/mL DMSO and selected DMSO as the solvent. ECBC and FAL selected the culture media as the solvent. The SMT selected DMSO as the solvent for acetaminophen to be used by all laboratories.

7.5 Summary

Intra- and inter-laboratory reproducibility were assessed using ANOVA, CV analysis, comparison of the laboratory-specific IC_{50}-LD_{50} regressions to one another (for each test method) and laboratory concordance for the GHS acute oral toxicity category predictions. ANOVA permits statistical comparisons of laboratories and experimental averages, while controlling for other factors. CV analysis is an empirical way of expressing the relative magnitudes of variability on a standardized scale. ANOVA results for the reference substances showed significant laboratory differences for 26 substances for the 3T3 NRU test method and seven substances for the NHK test method. Intralaboratory CV values were 1-122% for the 3T3 NRU test method and 1-129% for the NHK NRU test method. Mean intralaboratory CV values were 26% for both test methods, but the NHK NRU test method had a lower interlaboratory CV (28% vs 46%). Interlaboratory CV values were 2-135% for the 3T3 NRU test method and 1-99% for the NHK NRU test method. FAL had the highest mean intralaboratory CV for both test methods (33% for the 3T3 NRU test method and 42% for the NHK NRU test method).

An analysis to determine the relationship between the chemical attributes and interlaboratory CV indicated that physical form, solubility, and volatility had little effect on CV. CV seemed to be related, however, to chemical class, GHS acute toxicity category, IC_{50}, and boiling point. Reference substances in the amide class had unusually low mean interlaboratory CV values for both the 3T3 NRU test method (15%) and NHK NRU test method (16%) compared with the overall mean interlaboratory CV values (46% for the 3T3 NRU test method and 28% for the NHK NRU test method). Reference substances in the organophosphate and heterocyclic classes had unusually high mean interlaboratory CV values for the 3T3 NRU test method (74% and 71%, respectively), but not for the NHK NRU test method. Mean interlaboratory CV values were large for substances in the most toxic GHS acute categories, especially for the 3T3 NRU test method. The mean interlaboratory CV for substances in the classes for $LD_{50} \leq 5$ mg/kg (69%) and $5 < LD_{50} \leq 50$ mg/kg (78%) were larger than the mean overall interlaboratory CV (46%,) for the 3T3 NRU test method. For the NHK NRU test method, the mean interlaboratory CV was 37% for substances with $LD_{50} \leq 5$ mg/kg and 41% for substances with $5 < LD_{50} \leq 50$ mg/kg while the mean overall interlaboratory CV was 28%. A Spearman correlation analysis indicated that IC_{50} was negatively correlated to interlaboratory CV for both 3T3 ($p = 0.015$) and NHK ($p = 0.024$) NRU test methods and that boiling point was positively correlated to interlaboratory CV ($p = 0.021$) for the 3T3 NRU test method.

The analysis of interlaboratory reproducibility by evaluating the similarity of the laboratory specific IC_{50}-LD_{50} regressions indicated that the laboratory regressions for both test methods were not significantly different ($p < 0.05$) from one another ($p = 0.796$ for the 3T3 NRU and $p = 0.985$ for the NHK NRU). The evaluation of laboratory concordance for the prediction of GHS acute oral toxicity category when the laboratory-specific IC_{50} data were applied to the same regression yielded the following proportions of substances for which all laboratories agreed on the GHS acute oral toxicity categorization:

- 78% (52/67) for the 3T3 NRU and 87% (59/68) for the NHK NRU with the RC regression
- 81% (52/64) for the 3T3 NRU and 91% (59/65) for the NHK NRU with the RC rat only weight regression

- 84% for the both test methods (36/43 for the 3T3 NRU and 37/44 for the NHK NRU) with the RC rat only weight regression excluding substances with specific mechanisms of action

ANOVA results for the positive control, SLS, IC_{50} in the 3T3 NRU test method indicated that there were significant differences among laboratories ($p = 0.006$) and but not between study phases within laboratories ($p > 0.01$). However, interlaboratory CV values, which ranged from 2% to 10% for the study phases, indicated that the laboratories were similar. Intralaboratory CV values for the study phases ranged from 5% to 24%. SLS IC_{50} values for the NHK NRU test method were more variable than those for the 3T3 NRU test method. ANOVA results for SLS in the NHK NRU test method indicated that there were significant differences between laboratories ($p < 0.001$) and between study phases within laboratories ($p \leq 0.001$). A change in cell culture methods at FAL decreased the SLS IC_{50} from Phase Ib to Phase II. Intralaboratory CV values for the NHK NRU SLS IC_{50} during the various study phases ranged from 11% to 51%. Interlaboratory CV values for SLS in the NHK NRU test method ranged from 8% to 39%.

Cell culture medium was used as the solvent for testing 38 substances and DMSO was used for 34 substances. The laboratory concordance in selecting solvent for the reference substances using the solubility protocol was 76% (55/72).

[This Page Intentionally Left Blank]